

QUANTUM CHANNEL SIMULATION OF PHYLOGENETIC BRANCHING MODELS

Demosthenes Ellinas ^{*}, and Peter D Jarvis [†]

January 18, 2019

Abstract

Quantum channel simulations constructing probability tensors for biological multi-taxa in phylogenetics are proposed. These are given in terms of positive trace preserving maps (quantum channels), operating on quantum density matrices, using evolving systems of quantum walks with multiple walkers. Simulation of a variety of standard phylogenetic branching models, applying on trees of various topologies, is constructed using appropriate decoherent quantum circuits. For the sequences of biological characters so modelled, quantum simulations of statistical inference for them are constructed, given appropriate aligned molecular sequence data. This is achieved by the introduction of a quantum pruning map, operating on likelihood operator observables, utilizing state-observable duality and quantum measurement theory. More general strategies for related quantum simulation targets are also discussed.

1 Introduction

In the last two decades, quantum mechanics has found itself in a situation that could be characterized as an epistemological exodus. It has expanded its scope and applicability into new fields, such as information theory, the theory of computation, and even biology, and has addressed fundamental problems and procedures of these fields, by means of its physical-mathematical conceptual and computational apparatus [1, 2]. What were previously accepted as quantum paradoxes and oddities, like quantum entanglement, have turned out to be the keys to constructing novel computational and communicational algorithms, providing the means for launching a new quantum technology. In this vein, this paper puts forward a novel application of the discipline of quantum computation-information to the field of evolutionary phylogenetics [3, 4]. Phylogenetics' main task is to construct ancestral relationships (phylogenies), inferred by analyzing statistical data, collected for various (morphological or genotypic) kinds of characters or traits, possessed by selected groups of biological organisms (taxa). This amounts to construction of phylogenetic trees with appropriate branching patterns and evolutionary lengths, that successfully reproduce statistical trends of alignments of sequences of certain characters [3, 4]. Various evolutionary models that compete by adjusting their tree vertices and edge transition probabilities, to accomplish this computationally NP-hard task [5], are then assessed by some statistical inference technique such as maximum likelihood estimation (MLE) [3, 6].

In this work, quantum channel simulation [7] of molecular phylogenetic evolution and inference, is introduced in terms of trace preserving maps operating on quantum density matrices. In the DNA case, basic multi-parametric evolutionary models adopting standard choices of substitution rate matrices are

^{*}Technical University of Crete, School of Electrical and Computer Engineering QLab, 73 100 Chania Crete Greece (ellinas@ece.tuc.gr)

[†]Alexander von Humboldt Fellow. School of Natural Sciences (Mathematics and Physics), University of Tasmania, Tas 7001 Australia (peter.jarvis@utas.edu.au)

simulated, and an association between phylogenetic trees and quantum circuits is established. Specifically, standard models with uniform stationary base frequency are implemented with unitary channels and also via quantum walks, while the nonuniform case entails post-measurement state maps and also non-unitary channels. Finally, given aligned sequence data, a quantum protocol for the iterative pruning process for estimating maximum likelihood parameters for phylogenetic trees under simulation, is established in terms of measurements on appropriate quantum likelihood observables.

2 Gate operations for phylogenetic tensor simulation

Notation: Let the character set be $\Sigma = \{0, 1, \dots, N-1\} = \{0\} \cup \Sigma^*$. Here 0 is considered to be the “null” or no character symbol. Introduce the Hilbert space of character states $H \approx l_2(\Sigma) = \text{span}(|i\rangle; i \in \Sigma)$, of dimension $\dim H = |\Sigma|$, and consider the space $\text{Lin}(H)$ of linear operators acting on H . Examples are the complete set of projectors $\hat{P}_i = |i\rangle\langle i|$, $i \in \Sigma$, the shift operator $h|i\rangle = |i +_N 1\rangle$, with $+_N$ addition modulo N , (so that $h^N = 1$), and the space of density matrices $\mathcal{D}(H) \subset \text{Lin}(H)$. A classical (discrete) probability distribution is represented as a vector $(p_0, p_1, p_2, \dots, p_{N-1})$, and the corresponding quantum stochastic system is represented by a diagonal density matrix $\rho = \sum_{i \in \Sigma^*} p_i \hat{P}_i \in \mathcal{D}(H)$; for biological applications we will always assume $p_0 = 0$ (so that in practice the sum runs only over characters $i \in \Sigma^*$). On bipartite systems, the unitary control-not operator $U_{cn} \in \text{Lin}(H \otimes H)$ defined as $U_{cn} = \sum_{k \in \Sigma} \hat{P}_k \otimes h^k$, acts as [1] $U_{cn} |i\rangle \otimes |j\rangle = |i\rangle \otimes |j +_N i\rangle$.

2.1 Splitting, cladogenesis, speciation

The splitting operation [8, 9, 10] Δ for given 1-taxon matrix $\rho = \sum_{i \in \Sigma^*} p_i \hat{P}_i$, is implemented by the adjoint action of U_{cn}

$$\Delta\rho = U_{cn}(\rho \otimes \hat{P}_0)U_{cn}^\dagger = \sum_{i_1, i_2 \in \Sigma^*} p_{i_1 i_2} \hat{P}_{i_1} \otimes \hat{P}_{i_2}, \quad (1)$$

where $p_{i_1 i_2} = p_{i_1} \delta_{i_1 i_2}$, so $\Delta\rho$ is identified with a two-taxon density matrix. The control-not gate embedded in various positions in s -fold products of character spaces, for example $\mathbf{1}^{\otimes k-1} \otimes U_{cn} \otimes \mathbf{1}^{\otimes s-k-1}$ for speciation at position k , provides the way to construct s -taxon phylogenetic trees of various topologies [9]. By extension of the 1-taxon case, in order for U_{cn} to copy the contents of position k on to position $k+1$, in this iterative construction it is assumed that the target density operator entails a corresponding sum of the form $\sum p_{\dots i_k \dots} \dots \otimes \hat{P}_{i_k} \otimes \hat{P}_0 \otimes \dots$.

2.2 Phyletic evolution, anagenesis

For an s -taxon density matrix $\rho = \sum_{i_1, \dots, i_s \in \Sigma^*} p_{i_1 \dots i_s} \hat{P}_{i_1} \otimes \dots \otimes \hat{P}_{i_s}$, a suitable local unitary $U = \bigotimes_{i=1}^s U_i \in \text{Lin}(H)^{\otimes s}$, formalizes the phyletic evolution of taxa, when its action is composed with the s -fold product of the local *diagonalizing map* $\mathcal{E}_d^{\otimes s}$, where $\mathcal{E}_d(\cdot) = \sum_{k \in \Sigma} \hat{P}_k(\cdot)\hat{P}_k$, is the quantum channel, the completely positive trace preserving (CPTP) map that projects out the diagonal part of a matrix [11, 12, 13, 14], that is a decoherent map. Thus we have $\rho \rightarrow \tilde{\rho} \equiv \mathcal{E}_d^{\otimes s}(U\rho U^\dagger)$, where

$$\tilde{\rho} = \sum_{i_1, \dots, i_s \in \Sigma^*} \tilde{p}_{i_1 \dots i_s} \hat{P}_{i_1} \otimes \dots \otimes \hat{P}_{i_s}, \quad (2)$$

$$\tilde{p}_{i_1 \dots i_s} = \sum_{j_1, \dots, j_s \in \Sigma^*} p_{j_1 \dots j_s} (M_1 \otimes \dots \otimes M_s)_{i_1 j_1; \dots; i_s j_s}. \quad (3)$$

Abbreviating the adjoint action of an operator as $AdS(\cdot) \equiv S(\cdot)S^\dagger$, we say that the map $\mathcal{E}_d^{\otimes s}(AdU)$ thus induces a general doubly-stochastic transformation in the probability tensor. In (2) the Hadamard

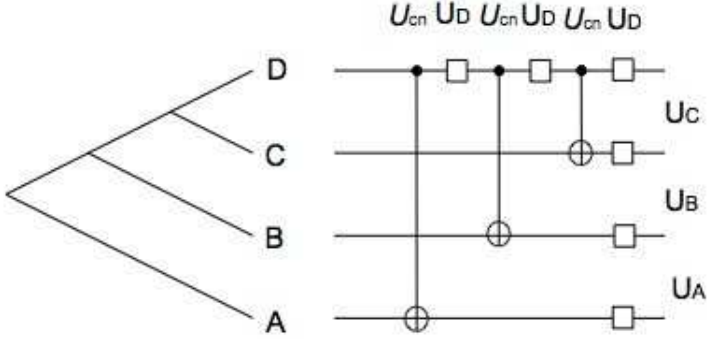


Figure 1: A 4-taxon tree and its simulating quantum circuit.

or entry-wise product of matrices defined [15] as $(A \circ B)_{ij} = A_{ij}B_{ij}$, has been used, to obtain the Markov matrices $M_i = U_i \circ U_i^*$, which will drive evolution on edges of a model phylogenetic tree. Below, we make particular choices of U to reflect different types of phylogenetic models, with some cases also requiring modifications of the channel and measurement protocols. Fig. 1 summarizes the preceding discussion, by showing a four taxon tree and its simulating quantum circuits.

2.3 Phyletic evolution and quantum walks

It has long been appreciated that faithful modeling of trait evolution in phylogenetics is problematic. As has been remarked, “... Brownian motion is a poor model, and so is Ornstein-Uhlenbeck” [16]¹.

As an example of the flexibility of the quantum simulation context, we here present a novel protocol for the stochastic phyletic evolution of traits, via quantum simulation employing the formalism of a quantum walk (QW) [18, 19, 20], operating locally on density matrices along edges of trees. This is set up as follows. Introduce, in addition to character Hilbert space H (the “walker” space), at each node of phylogenetic tree an auxiliary “coin” Hilbert space $H_c \approx l_2(C) = \text{span}(|+\rangle, |-\rangle)$, and projectors $P_{\pm} \in \text{Lin}(H_c)$. Evolution now proceeds on joint “walker” and “coin” states $\rho_c \otimes \rho$ via a standard QW conditional unitary operator $V = (P_+ \otimes h + P_- \otimes h^\dagger)U \otimes \mathbf{1}$, acting from $H_c \otimes H$ to itself. One “step” of such a QW is realized by the map on the “walker” density matrix, viz. $\rho \rightarrow \mathcal{E}_{V^k}(\rho) := \text{Tr}_c V^k(\rho_c \otimes \rho) V^{\dagger k}$, followed by diagonalization with \mathcal{E}_d . For s taxa, $E_{V^k} \equiv (\mathcal{E}_d^{\otimes s} \circ \mathcal{E}_{V^2}^{\otimes s})$.

For example, for the two-taxon case, with $k = 2$ and coin initially in a pure state $\rho_c = |c\rangle\langle c|$ with $|c\rangle = |+\rangle$ or $|-\rangle$, we obtain $\rho \rightarrow \tilde{\rho} = E_{V^2}(\rho)$ with $\tilde{\rho} = \sum_{mn} \tilde{p}_{mn} \hat{P}_m \otimes \hat{P}_n$ where $\tilde{p}_{mn} = \sum_{ab} p_{m-a, n-b} q_a^{(c)} q_b^{(c)}$, and $q_a^{(c)} := \sum_d M_{d,a-d} M_{d-a,c}$ determined by the coin tossing unitary U via the Hadamard product $M = U \circ U^*$. By re-labelling of indices mod N , the map $p \rightarrow \tilde{p}$ can be identified in this case with independent evolution on each edge, via the substitution matrix $Q_{a,m} := q_{m-a}^{(c)}$ associated with the probability distribution $q_a^{(c)} > 0, \sum_a q_a^{(c)} = 1$ ². More generally however, stochastic tensors such as \tilde{p} and their multi-taxon generalizations, as objects of quantum simulations, are natural candidates for alternative models, of potential interest in view of the above-mentioned foundational difficulties of handling trait evolution in phylogenetics.

Finally, we note that here (and in the examples below) the diagonalizing map \mathcal{E}_d itself can be cast in the form of a CPTP map, i.e. $\mathcal{E}_d(\rho) = \sum_{k \in \Sigma^*} \hat{P}_k \rho \hat{P}_k = \sum_{k \in \Sigma^*} q_k U_k \rho U_k^\dagger$ with each $q_k = 1/|\Sigma|$, thanks

¹For related discussion see the video postings [17].

²A standard Markov model of the group-like class, with symmetry \mathbb{Z}_N [3, 4]

to the non-uniqueness of the operator sum representation, with unitaries U_k related to projectors by discrete Fourier transform, $U_k = \sum_l \omega^{kl} \hat{P}_l$ and $\omega = \exp(i2\pi/|\Sigma|)$. Below, similar quantum prescriptions will be given to the structural maps of standard evolutionary models.

3 Evolutionary models and quantum maps

Next we exploit the above considerations of phylogenetic models for DNA ($|\Sigma^*| = 4$) by exemplifying the quantum channel simulations for some standard parametrizations of the edge substitution matrix.

3.1 Uniform stationary distribution

The following models: Jukes-Cantor (JC), Kimura two-parameter (K2), Kimura three-parameter (K3), and the binary symmetric model (B), all have a uniform stationary distribution [21, 22, 23]. Firstly, we give in each case (denoted τ below) a direct Kraus representation of the composite quantum map $E_\tau \equiv \mathcal{E}_d \circ \mathcal{E}_\tau$. This is followed by a QW formulation using, as above, an additional ancillary “coin” space. Let X, Z denote the usual single qubit not and phase gates (the Pauli matrices σ_x, σ_z respectively) and $U_{kl} = X^k \otimes X^l$, for $k, l = 0, 1$. The following propositions are verified by direct calculation for operators in $l_2(\Sigma^*)$ acting on $\rho = \sum_{m \in \Sigma^*} p_m \hat{P}_m$:

Proposition K: Let $|\Sigma^*| = 4$ and $\tau \in \{K3, K2, JC\}$. We have

$$\begin{aligned} E_\tau(\rho) &= \sum_{k,l} \lambda_{kl}^{(\tau)} U_{kl}(\rho) U_{kl}^\dagger = \sum_{m \in \Sigma^*} (M_\tau p)_m \hat{P}_m, \\ M_\tau(a, b, c) &= \sum_{kl} \lambda_{kl}^{(\tau)} U_{kl} \circ U_{kl}^* = \sum_{kl} \lambda_{kl}^{(\tau)} X^k \otimes X^l. \end{aligned} \quad (4)$$

The weights $\lambda_{kl}^{(\tau)}$ and corresponding model Markov matrices M_τ are defined as follows. For generic parameters define the weights $\lambda_{kl}(a, b, c)$ as $\lambda_{00} = 1 - a - b - c$, $\lambda_{10} = a$, $\lambda_{01} = b$, $\lambda_{11} = c$, and take the corresponding convex sum $M(a, b, c)$. Then $\lambda_{kl}^{(3K)} = \lambda_{kl}(a, b, c)$, $M_{3K} \equiv M(a, b, c)$, $\lambda_{kl}^{(2K)} = \lambda_{kl}(a, b, b)$, $M_{2K} \equiv M(a, b, b)$, and finally $\lambda_{kl}^{(JC)} = \lambda_{kl}(a, a, a)$, $M_{JC} \equiv M(a, a, a)$. \square

Proposition K': The CPTP map E_τ has, in addition to the operator sum representation above, also a QW like representation $E_\tau(\rho) = \text{Tr}_c V_\tau(\rho_c \otimes \rho) V_\tau^\dagger$, in terms of a unitary dilation $V_\tau = (\sum_{kl} P_k \otimes P_l \otimes U_{kl}) U_\tau \otimes \mathbf{1}$ which acts on a composite coin-walker space $H_c \otimes H$, with four-dimensional ancillary space. Here V_τ is a control-control- U_{kl} operator. For a coin density matrix with spectral decomposition $\rho_c = \sum_k \mu_k |c_k\rangle \langle c_k|$, the coin-tossing unitary U_τ should satisfy $\langle kl | U_\tau \circ U_\tau^* | c \rangle = \lambda_{kl}^{(\tau)}$, with $|c\rangle = \sum_k \mu_k |c_k\rangle$ a stochastic vector. Also $U_{kl} = e^{i\mathcal{H}_{kl}}$ where $\mathcal{H}_{kl} = \frac{1}{2}\pi[-(k+l)\mathbf{1} \otimes \mathbf{1} + kX \otimes \mathbf{1} + l\mathbf{1} \otimes X]$. \square

Proposition B: Let $|\Sigma^*| = 2$. The map $\mathcal{E}_d \circ E_B$, where $E_B(\rho) = (1-a)\rho + aX\rho X^\dagger$, simulates the binary symmetric model $M_B(a) = (1-a)\mathbf{1} + aX$ acting as $\rho = \sum_{m \in \Sigma^*} p_m \hat{P}_m \rightarrow \sum_{m \in \Sigma^*} (M_B p)_m \hat{P}_m$. \square

Proposition B': The “control flip” map E_B is unitarized in composite coin-walker space with a two-dimensional ancillary space as, $E_B(\rho) = \text{Tr}_c V_B(\rho_c \otimes \rho) V_B^\dagger$, with the starting coin state $\rho_c = |1\rangle \langle 1|$, and $V_B = \sqrt{a}\mathbf{1} \otimes \mathbf{1} + \sqrt{1-a}Y \otimes X$, and $Y \equiv ZX$. \square

Remark: In the QW picture, the weight parameters $\lambda_{kl}^{(\tau)}$ determine non-uniquely, via the unistochastic [24] matrix $U_\tau \circ U_\tau^*$, the coin-tossing matrix U_c , which in turn determines the U -quantization of the underlying classical walk [25] with evolution matrix $V_{cl} \equiv \sum_{kl} P_k \otimes P_l \otimes U_{kl}$.

3.2 Non-uniform stationary distribution

We treat the Felsenstein model [6] (F), and the strand symmetric model [26, 27] (SSM), as illustrative of models with non-uniform stationary distributions. For the Felsenstein model, quantum simulation requires the following ingredients. The model's stationary distribution $(\pi_1, \pi_2, \pi_3, \pi_4)$, $\sum \pi_i = 1$, is to be used to introduce the observable $\mathbf{1}_\pi := 4 \sum_i \pi_i \hat{P}_i$, with Kraus operators $F_{ij} = \sqrt{\pi_j} |i\rangle \langle j|$, $i, j \in \Sigma = \{1, 2, 3, 4\}$ obeying the resolution relation $\sum_{ij} F_{ij}^\dagger F_{ij} = \frac{1}{4} \mathbf{1}_\pi$. Again let $\rho = \sum_{m \in \Sigma^*} p_m \hat{P}_m$. By direct calculation we obtain:

Proposition F: The quantum map implementing the Felsenstein model $\rho \rightarrow E_F(\rho) = \sum_{m \in \Sigma^*} (M_F p)_m \hat{P}_m$ is given by

$$E_F(\rho) = (1 - a) \frac{1}{p_\pi} \sum_{i,j} F_{ij} \rho F_{ij}^\dagger + a \rho, \quad (5)$$

where $p_\pi = \text{Tr}(\sum_{i,j} F_{ij}^\dagger F_{ij} \rho) = \text{Tr}(\frac{1}{4} \mathbf{1}_\pi \rho)$ is a normalization constant, and the model's stochastic matrix is obtained as $M_F = (1 - a) \sum_{i,j} F_{ij} \circ F_{ij} + a \mathbf{1}$. \square

In the framework of quantum measurement theory, simulation of the Felsenstein model is interpreted as follows. There are two observables: $\mathbf{1}_\pi$ as above, and also $\mathbf{1}_\pi^\#$ defined analogously in terms of the complementary probability distribution $(\pi_1^\#, \pi_2^\#, \pi_3^\#, \pi_4^\#)$, with $\mathbf{1} = \mathbf{1}_\pi + \mathbf{1}_\pi^\#$ forming a non-orthogonal decomposition of unity. These observables are measured by means of the so called instruments [11], which are the two families of Kraus generators: the $\{F_{ij}\}_{i,j=0}^1$ as above, and the analogous $\{F_{ij}^\#\}_{i,j=0}^1$ defined in terms of $\pi^\#$ rather than π (see for example [11]). The measurement probabilities of the observables $\mathbf{1}_\pi$ and $\mathbf{1}_\pi^\#$ in the system are $p_\pi^\# = \text{Tr}(\mathbf{1}_\pi^\# \rho)$, and $p_\pi = \text{Tr}(\mathbf{1}_\pi \rho)$, and the action of quantum map E_F on the density matrix $E_F(\rho)$ gives the post-measurement density matrix for a *non-efficient* quantum measurement for observable $\mathbf{1}_\pi$ of *finite strength* [28]. The complementary measurement of $\mathbf{1}_\pi^\#$ is not used. In the uniform limit $\pi_j = \frac{1}{4}$ then $\mathbf{1}_\pi = \mathbf{1}$, $\mathbf{1}_\pi^\# = \mathbf{0}$ and $p_\pi = 1$, and the model reduces to the JC model.

As a final example, we consider the quantum channel presentation of the strand symmetric model [26, 27].

Proposition SSM: The quantum map $\rho \rightarrow E_M(\rho)$ operating on a 4×4 density matrix ρ explicitly given by

$$E_M(\rho) = (\sqrt{\lambda} P_0 + \sqrt{\mu} P_1) \otimes \mathbf{1} \rho ((\sqrt{\lambda} P_0 + \sqrt{\mu} P_1) \otimes \mathbf{1})^\dagger + V_1 \otimes \mathbf{1} \rho (V_1 \otimes \mathbf{1})^\dagger + V_2 \otimes \sigma_1 \rho (V_2 \otimes \sigma_1)^\dagger, \\ V_1 = \begin{pmatrix} 0 & \sqrt{e} \\ \sqrt{b} & 0 \end{pmatrix}, \quad V_2 = \begin{pmatrix} \sqrt{a} & \sqrt{f} \\ \sqrt{c} & \sqrt{g} \end{pmatrix},$$

with parameters $\lambda = 1 - (a + b + c)$, $\mu = 1 - (f + g + e)$, provides a non-unitary channel implementation of the strand symmetric phylogenetic substitution model. Again, assuming the initial density matrix is diagonal, $\rho = \sum p_i \hat{P}_i$, the map $\rho \rightarrow E_M(\rho)$ implements the stochastic transformation $p \rightarrow M \cdot p$ under $M = (\sqrt{\lambda} P_0 + \sqrt{\mu} P_1) \circ (\sqrt{\lambda} P_0 + \sqrt{\mu} P_1)^* + V_1 \circ V_1^* + V_2 \circ V_2^*$ which is the SSM substitution [27]. \square

Remark: While the above quantum implementations of models with uniform stationary distributions (and symmetric substitution matrices) have used unitary channel generators, in the non-uniform cases, extensions of the formalism have been exploited: in the Felsenstein case, via non-efficient measurement protocols, and in the strand symmetric case, via nonunitary channels. It can be expected that for other classes of models, for example, in the DNA case, substitution matrices in the so-called Lie-Markov hierarchy [29, 30, 31, 32], as well as the general Markov model, similar considerations will apply. The present examples serve as illustrations of the general constructions required.

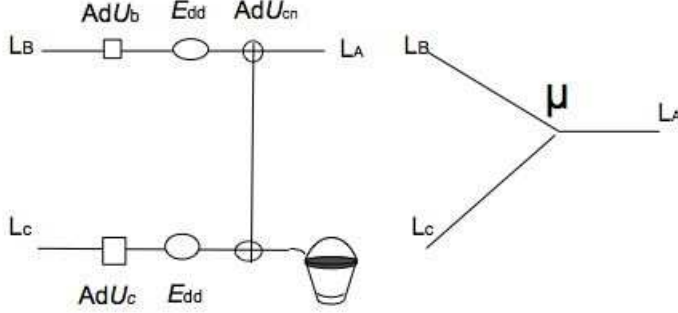


Figure 2: Circuit for pruning map of likelihood operators.

4 Quantum estimation of likelihood

Our general framework also encompasses the quantum estimation of model-based tree likelihood parameters [6], whose numerical calculation and optimization provides a major tool for phylogenetic inference (computational heuristics are discussed in [33]; see for example [34] for a recent re-analysis). Likelihood evaluation (incorporating tree searching) has been demonstrated to be a computationally NP-hard problem [5, 6], and it is therefore desirable to put forward an equivalent quantum simulation realization. In the usual formulation [6], likelihood vectors are initialized at the pendant nodes (leaves) of a tree, and are then computed recursively back to the root node, the final result being a scalar quantity, the tree likelihood. The key operation is that of pruning, that is, of arriving at the likelihood for a parent node, say A , by combining a pair of daughter likelihoods, say B, C , from nodes which root two sub-trees. Explicitly, likelihoods for daughter nodes B, C are combined to give the parent likelihood $L_i^A = (\sum_{j \in \Sigma^*} M_{ij}^B(t_B) L_j^B) (\sum_{k \in \Sigma^*} M_{ik}^C(t_C) L_k^C)$, where $M^{B,C}(t_{B,C})$ are stochastic matrices depending on branch lengths $t_{B,C}$ specified by the evolutionary model employed. Next, an alignment of s taxa over Λ sites is considered. If the characters at site l of the alignment are $i_1^{(l)}, i_2^{(l)} \dots i_s^{(l)}$, then likelihoods for the tips of the tree (leaf nodes) are initialized to $L_k^{(l)} = \delta(k, i_k^{(l)})$. The pruning map is applied recursively at all cherries, and then higher up the tree, to arrive at the total tree likelihood $L_{tr}^{(l)} = (L_{tr\ k}^{(l)})_{k=1}^s$, which is finally averaged over the assumed stationary distribution (π_i) of the model to obtain site l 's likelihood $L^{(l)} = \sum_k \pi_i L_{tr\ k}^{(l)}$. For the entire alignment, the tree (log) likelihood is then $L(T; w^*) = \max_w \log \Pi_{l=1}^\Lambda L^{(l)}$, where T denotes the tree topology and w^* the optimal model (weight) parameters.

In the quantum simulation introduced here, likelihoods are regarded as quantum observables, that is operators in $Lin(H)$, dual to density operators under the trace inner product (see above). The likelihood operator at node A has components $\hat{L}_i^A \equiv L^A(t|i) = \mathbb{P}(i|t)$, where $\mathbb{P}(i|t)$ is the conditional probability of character $i \in \Sigma^*$, for parameters $t = (T, w)$. Here $A = 1, 2, \dots, s$ are leaf nodes and $A = s + 1, \dots, 2s - 2$, internal (ancestral) nodes. Consider parent and daughter nodes A, B and C , with respective likelihood operators \hat{L}^A, \hat{L}^B and \hat{L}^C . Operators for daughter nodes B, C are combined using the analog of pruning, the quantum pruning map $\mu : Lin(H) \otimes Lin(H) \rightarrow Lin(H)$ that provides the parent operator $\hat{L}^A = \mu(\hat{L}^B \otimes \hat{L}^C)$, where $\mu = Tr_B \circ Ad U_{cn}^\dagger \circ \mathcal{E}_{dd} \circ Ad(U_B \otimes U_C)$. The map μ uses stochastic matrices $M^x(t_x) = U_x \circ U_x^*$, depending on branch lengths t_x for $x = A, B$, as given by the model employed, and the collective “diagonalizing map” $\mathcal{E}_{dd}(\cdot) = \sum_k \hat{P}_k \otimes \hat{P}_k(\cdot) \hat{P}_k \otimes \hat{P}_k$. Fig. 2 presents a quantum circuit realizing map μ . By using its embedding $\mu_{r,r+1} = id^{\otimes r-1} \otimes \mu \otimes id^{\otimes s-r}$ for various values of r according to the topology of the binary tree, the pruning map μ is applied recursively to all cherries, and then higher up the tree. In this way we arrive at the tree likelihood operator $\hat{L}_{tr}^{(l)}$, which then is contracted with

model's stationary density matrix $\rho^\pi = \sum_i \pi_i \hat{P}_i$, to yield as a measurement result the site l likelihood $L^{(l)} = \text{Tr}(\hat{L}_{tr}^{(l)} \rho^\pi) \equiv \langle \hat{L}_{tr}^{(l)}, \rho^\pi \rangle$. For the entire alignment, the tree (log) likelihood is (c.f. the identity $\text{Tr}(AB)\text{Tr}(CD) = \text{Tr}(A \otimes C)(B \otimes D)$)

$$L = \max_w \log \prod_{l=1}^{\Lambda} \langle \hat{L}_{tr}^{(l)}, \rho^\pi \rangle = \max_w \log \text{Tr}(\otimes_{l=1}^{\Lambda} \hat{L}_{tr}^{(l)} \rho_\Lambda),$$

where $\rho_\Lambda \equiv (\rho^\pi)^{\otimes \Lambda}$ is the product of Λ stationary density matrices.

In fact this Heisenberg-like picture of updating the observables (likelihoods), and finally contraction with the stationary density matrix to derive site and eventually alignment likelihoods, can be converted to a Schrödinger-like picture, using the observable-state duality, exemplified here by the trace cyclic property. Firstly note that the pruning map can be expressed as $\mu(\hat{L}^B \otimes \hat{L}^C) = \nu^{-1} \mathcal{E}_B(\hat{L}^C)$, $\nu = \text{Tr} \hat{L}^B$, where the positive stochastic map \mathcal{E}_B decomposes as $\mathcal{E}_B \equiv \mathcal{E}_{Bpd} \circ \text{Ad } U_C$ with $\mathcal{E}_{Bpd}(\cdot) = \sum_{k \in \Sigma} q_k^B \text{Ad } \hat{P}_k(\cdot)$, a probabilistic diagonalizing map, with probabilities $q_k^B = \nu^{-1} \langle k | U_B \hat{L}^B U_B^\dagger | k \rangle$. As the roles of \hat{L}^B and \hat{L}^C can be exchanged above with appropriate modification, (\mathcal{E}_B becomes \mathcal{E}_C etc), we note that μ is proportional to a stochastic map either way, and by duality it can be made to act on density matrices instead of likelihood operators. This is also true for embedded pruning maps $\mu_{r,r+1}$, that is, they will also be proportional to maps $\mathcal{E}_{B;r,r+1}$ for the appropriate current likelihood \hat{L}^B etc. Then the tree likelihood operator $\hat{L}_{tr}^{(l)}$, obtained by composing pruning maps, will eventually be described by pruning a final cherry, say with nodes B_f and C_f , ie. $\hat{L}_{tr}^{(l)} = \nu_f^{-1} \mathcal{E}_{B_f}(\hat{L}^{C_f})$, $\nu_f = \text{Tr} \hat{L}^{B_f}$. Then the likelihood at site l is obtained as $L^{(l)} = \nu_f^{-1} \text{Tr}(\mathcal{E}_{B_f}(\hat{L}^{C_f}) \rho^\pi) = \nu_f^{-1} \text{Tr}(\hat{L}^{C_f} \mathcal{E}_{B_f}^*(\rho^\pi))$, where the dual map $\mathcal{E}_{B_f}^*$ of \mathcal{E}_{B_f} acting on the density matrix is introduced. This situation is extended similarly to the likelihood of the entire alignment by assigning additional site indices l to each likelihood operator, e.g. $\hat{L}_l^{B_f}$ and $\hat{L}_l^{C_f}$, as well as trace coefficients ν_f^l etc, to obtain $L(T; w^*) = \max_w \log \prod_{l=1}^{\Lambda} (\nu_f^l)^{-1} \text{Tr}(\hat{L}_\Lambda(\otimes_{l=1}^{\Lambda} \mathcal{E}_{B_f,l}^*) \rho_\Lambda)$. Here $\hat{L}_\Lambda \equiv \otimes_{l=1}^{\Lambda} (\hat{L}_l^{C_f})$ is the product of Λ different likelihood operators, corresponding to final cherries of the respective trees, employed to construct tree likelihoods. Note that $\otimes_{l=1}^{\Lambda} \mathcal{E}_{B_f,l}^*$ is a collective factorized map that can be expressed in terms of a unitary dilation, and this would in principle be implemented by a Hamiltonian quantum model.

A maximum likelihood estimation of a single unknown parameter determining a quantum walk map that would in turn be used to induce phyletic evolution, along a branch of a phylogenetic tree, as has been previously outlined, is indeed possible; such an explicit calculation has been carried out for a typical QW [35]. The MLE problem could be generalized for phyletic evolution parameters (branch lengths) for a whole tree along the lines of that work. The operational character that such an MLE can be given, that is its reformulation in the language of quantum evolution and quantum measurements, is also explicitly spelt out. Also important is the fact that in this quantum algorithm MLE problem, the data availability problem for safe estimation, here in the guise of the well known question “more taxa or more characters” (see for example [36] and references therein), shows up in the quantum context as well. The well known quadratic diffusion speed up of a QW compared to a classical walk [18], can in fact be utilized to achieve gathering more data with fewer QW steps, than in the classical MLE based on a classical random walk [35].

In view of the results of the present work, and in combination with those in [35], we would claim that the simulation of phylogenetic evolutionary models as well as the problem of parameter estimation, could be considered as a form of quantum computation.

5 Conclusions and Prospects

In conclusion, this study lays the groundwork for simulating, by quantum mechanical means [37]-[43], the probability tensors of multi-taxa systems, and for estimating the maximal likelihood of a phylogenetic

alignment. One general remark concerning the general strategy adopted here for the quantum simulation of phylogenetics states and processes would refer to the choice of introducing the local *diagonalizing map* \mathcal{E}_d that zeroes the off diagonal elements (coherences) of the processed density matrix. This choice obviously renders the quantum simulation able to reproduce the results, namely the probability tensors, of the classical phylogenetic evolutionary models. This in fact was the main target of the paper. At the same time, this choice seems to be a very specific one, in particular due to the fact it makes no use of the quantum coherences developed step-wise during the implementation of quantum circuits that achieve the classical results. Obviously there are more general simulation strategies that would avoid diagonalizing the density matrices, and instead exploit quantum coherence in order to explore its possible utility in the context of phylogenetic evolutionary models, the more so given that coherence is recently treated as quantum resource (see for example [44], and references therein). Investigations of the role of the coherence resource in the context of quantum algorithms and quantum communications form an established research area. The present work thus serves also as a position paper regarding the feasibility of the simulation of phylogenetic modelling by quantum processes. Addressing quantum aspects of simulated phylogenetic models would be an exciting new area, for which the present paper, in terms of its set scope and results, may constitute a prerequisite step.

Also as an alternative to the choice of encoding and manipulating phylogenetic probabilistic information on the diagonal elements of density matrices, we should mention that the spectra, that is the eigenvalues of density matrices regarded as probability distributions, offer an alternative framework for quantum simulation of phylogenetics processes. One can actually work out (bi)stochastic matrices connecting the spectrum of input and output density matrices transformed by various quantum channels, and formulate a simulation framework of phylogenetic models based on the spectral transformations of density operators. Comparative advantages of eigenvalues versus diagonal elements in simulation would thus be a topic for further work.

Finally, in a more specific context, we would remark that with the tools developed here, prominent among problems for future investigation would be for example, a quantum computational simulation analysis of Steel’s conjecture regarding the complexity of tree search and parameter recovery in molecular phylogenetics modelling, and its resolution [45, 46].

Acknowledgements

PDJ thanks the Technical University of Crete, School of ECE QLab, for hospitality during collaborative visits. The Australian-American Fulbright Foundation, and staff and colleagues at the Department of Statistics, University of California Berkeley, and the Department of Physics, University of Texas Austin, for visits as an Australian senior Fulbright scholar, are also acknowledged. Part of this work was carried out under an ARC discovery project.

References

- [1] M. A. Nielsen and I. L. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, 2000.
- [2] Markus Arndt, Thomas Juffmann, and Vlatko Vedral. Quantum physics meets biology. *HFSP journal*, 3(6):386–400, 2009.
- [3] J. Felsenstein. *Inferring Phylogenies*. Sinauer Associates, Sunderland, 2004.
- [4] Mike Steel. *Phylogeny: Discrete and random processes in evolution*. SIAM, Philadelphia, 2016.
- [5] Benny Chor and Tamir Tuller. Maximum likelihood of evolutionary trees is hard. In *Annual International Conference on Research in Computational Molecular Biology*, pages 296–310. Springer, 2005.
- [6] J. Felsenstein. Evolutionary trees from DNA sequences: a maximum likelihood approach. *Journal of Molecular Evolution*, 17:368–376, 1981.

- [7] Benjamin Schumacher. Sending entanglement through noisy quantum channels. *Physical Review A*, 54(4):2614, 1996.
- [8] P. D. Jarvis and J. D. Bashford. Quantum field theory and phylogenetic branching. *Journal of Physics A: Mathematical and General*, 34:L703–L707, 2001.
- [9] J. G. Sumner, M. A. Charleston, L. S. Jermin, and P. D. Jarvis. Markov invariants, plethysms, and phylogenetics. *J. Theor. Biol.*, 253:601–615, 2008.
- [10] J. G. Sumner, B. H. Holland, and P. D. Jarvis. The algebra of the general Markov model on trees and networks. *Bull. Math. Biol.*, 74, 2012.
- [11] E Brian Davies and John T Lewis. An operational approach to quantum probability. *Communications in Mathematical Physics*, 17(3):239–260, 1970.
- [12] Karl Kraus. *States, effects and operations: fundamental notions of quantum theory*, volume 190 of *Lecture Notes in Physics*. Springer, 1983.
- [13] Masanao Ozawa. Quantum measuring processes of continuous observables. *Journal of Mathematical Physics*, 25(1):79–87, 1984.
- [14] John Watrous. *The theory of quantum information*. Cambridge University Press, 2018.
- [15] Albert W Marshall, Ingram Olkin, and Barry C Arnold. *Inequalities: theory of majorization and its applications*, volume 143. Springer, 1979.
- [16] J. Felsenstein. Multiple regressions with continuous and categorical data. *R-sig-phylo* mailing list, <https://stat.ethz.ch/pipermail/r-sig-phylo/2008-April/000053.html>, 2008.
- [17] J. Felsenstein. On Brownian motion, phylogenies and quantitative genetics. (Video postings <https://www.youtube.com/watch?v=cvhmC9i6YsU> and <https://www.youtube.com/watch?v=0RRvrlrWng>). 2014.
- [18] Julia Kempe. Quantum random walks: an introductory overview. *Contemporary Physics*, 44(4):307–327, 2003.
- [19] Dorit Aharonov, Andris Ambainis, Julia Kempe, and Umesh Vazirani. Quantum walks on graphs. In *Proceedings of the thirty-third annual ACM symposium on Theory of computing*, pages 50–59. ACM, 2001.
- [20] Andris Ambainis, Eric Bach, Ashwin Nayak, Ashvin Vishwanath, and John Watrous. One-dimensional quantum walks. In *Proceedings of the thirty-third annual ACM symposium on Theory of computing*, pages 37–49. ACM, 2001.
- [21] T. H. Jukes and C. R. Cantor. *Mammalian protein metabolism.*, chapter Evolution of Protein Molecules, pages 21–132. New York: Academic Press, 1969.
- [22] Motoo Kimura. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution*, 16(2):111–120, 1980.
- [23] Motoo Kimura. Estimation of evolutionary distances between homologous nucleotide sequences. *Proceedings of the National Academy of Sciences*, 78(1):454–458, 1981.
- [24] Ingemar Bengtsson, Åsa Ericsson, Marek Kuś, Wojciech Tadej, and Karol Życzkowski. Birkhoff’s polytope and unistochastic matrices, $n=3$ and $n=4$. *Communications in Mathematical Physics*, 259(2):307–324, 2005.
- [25] Demosthenes Ellinas and Ioannis Smyrnakis. Asymptotics of a quantum random walk driven by an optical cavity. *Journal of Optics B: Quantum and Semiclassical Optics*, 7(7):S152, 2005.
- [26] M. Casanellas and S. Sullivant. *Algebraic Statistics for Computational Biology*, chapter The Strand Symmetric Model, pages 305–321. Cambridge University Press, New York, 2005.
- [27] Peter D. Jarvis and Jeremy G. Sumner. Matrix group structure and Markov invariants in the strand symmetric phylogenetic substitution model. *Journal of Mathematical Biology*, 73(2):259–282, 2016.

- [28] Christopher A Fuchs and Kurt Jacobs. Information-tradeoff relations for finite-strength quantum measurements. *Physical Review A*, 63(6):062305, 2001.
- [29] J. G. Sumner, J. Fernández-Sánchez, and P. D. Jarvis. Lie Markov models. *Journal of Theoretical Biology*, 298:16–31, 2012.
- [30] Jesús Fernández-Sánchez, Jeremy G Sumner, Peter D Jarvis, and Michael D Woodhams. Lie Markov models with purine/pyrimidine symmetry. *Journal of Mathematical Biology*, 70(4):855–891, 2015.
- [31] Michael D. Woodhams, Jesús Fernández-Sánchez, and Jeremy G. Sumner. A new hierarchy of phylogenetic models consistent with heterogeneous substitution rates. *Systematic Biology*, 64(4):638–650, 2015.
- [32] P D Jarvis and J G Sumner. Systematics and symmetry in molecular phylogenetic modelling: perspectives from physics. ([arxiv:1809.03078v2](https://arxiv.org/abs/1809.03078v2) [q-bio]) Topical review, 2018.
- [33] A Stamatakis. RAxML-VI-HPC: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics*, 22:2688–2690, 2006.
- [34] J. G. Sumner and M. A. Charleston. Phylogenetic estimation with partial likelihood tensors. *J Theor Biol.*, 262:413–24, 2010.
- [35] D. Ellinas and P. D. Jarvis. Maximum likelihood parameter estimation via reluctant discrete quantum walk. (in preparation), 2018.
- [36] Andrew Mitchell, Charles Mitter, and Jerome C Regier. More taxa or more characters revisited: combining data from nuclear protein-encoding genes for phylogenetic analyses of noctuoidea (insecta: Lepidoptera). *Systematic Biology*, 49(2):202–224, 2000.
- [37] Feynman, R. Simulating Physics with Computers. *Int. J. Theor. Phys.* 21467–488, 1982
- [38] Lloyd, S. Universal Quantum Simulators. *Science* 2731073–1078, 1996
- [39] Somma, R., Ortiz, G., Gubernatis, J.E., Knill, E., Laflamme, R. Simulating physical phenomena by quantum networks. *Phys. Rev. A* 65042323, 2002
- [40] Buluta, I., Nori, F. Quantum Simulators. *Science* 326108–111, 2009
- [41] Cirac, J. I, Zoller P. Goals and opportunities in quantum simulation. *Nature Physics* 8(4)264–266, 2012
- [42] Johnson, T. H., Clark, S. R., Jaksch, D. What is a quantum simulator?. *EPJ Quantum Technology* 1(10), 2014
- [43] D. G. Angelakis Ed. *Quantum Simulations with Photons and Polaritons* Springer, 2017
- [44] Alexander Streltsov, Gerardo Adesso, and Martin B Plenio. Colloquium: quantum coherence as a resource. *Reviews of Modern Physics*, 89(4):041003, 2017.
- [45] M. Steel. <http://www.math.canterbury.ac.nz/~m.steel/files/misc/conjecture.pdf>. 2001.
- [46] Constantinos Daskalakis, Elchanan Mossel, and Sébastien Roch. Evolutionary trees and the Ising model on the Bethe lattice: a proof of Steel’s conjecture. *Probability Theory and Related Fields*, 149(1-2):149–189, 2011.